

The Bioinformatics Lab - Practical Course



Matúš Kalaš
Timothy Karl
Dr. Laszlo Kajan
Dr. Lothar Richter



Tuesdays 15:00-18:00 practical room MI 01.08.021

Guest Course Graduates (now PhD candidates)

- Maximilian Hecht



- Tatyana Goldberg



- What is this course like as a student?
- What did you like most and least in the course?
- What have you used, and foresee to use that you have learnt on this course?

Laszlo Kajan

- MSc: own RedHat Linux desktop
- PhD: small RedHat cluster made of desktops
- PostDoc (Poland): medium Linux cluster of desktops, eventually Debian Stable
- PostDoc (New York): RHEL cluster on rack
- PostDoc (Munich): Debian Stable cluster: initial architect and administrator with Guy Yachdav



Laszlo Kajan

- Open Source system administrator (600+ cores, 100+TB storage)
- Grid, virtualization and cloud experience
- Perl, C/C++, PHP, MySQL, JavaScript, some Python
- Bison, flex
- XML/XSD, some WSDL
- Debian and RPM packaging
- Debian Developer since 2013
 - 35 packages in Debian and derivatives (e.g. Ubuntu, Bio-Linux, CloudBioLinux)
 - PredictProtein, HHblits



Round of Introduction

- Do you have experience with free software?
 - With anything free for bioinformatics?
- Have you ever written a program others (also) wanted to use?

The Bioinformatics Lab - Objectives

- Learn to build an open source bioinformatics lab
- Get familiar with available open source software solutions
- Work more efficiently knowing about the available tools and what is in the background
- Acquire skills to contribute back your own work
- Make a contribution by taking part in the *semester challenge*



Organisation – Practical Session

- 30': homework (aka. programming challenge) presentation and discussion
 - one student
 - some slides
 - presentation to be published on course Wiki page
<https://rostlab.org/owiki/index.php/Tbl2013>
- 40': presentation and discussion of new topics
 - one student
 - presentation interleaved with questions and discussion
 - students to ask at least one question
- 20': general discussion and outline of new programming challenge
- Remaining: work on programming challenge

Organisation – Individual Work

- Completely implement programming challenge
 - there should be a working and documented solution on the virtual machine
- Prepare a presentation when it is your turn
 - some slides
 - published on course Wiki
- Prepare for the new topics
 - prepare to ask questions
 - useful links are provided on course wiki page

Requirements to Pass

- One presentation of homework
- One presentation of new topics
- At least one question asked per session
- Pass vim test and command line test
- Successfully prepare your share of the *semester challenge*
- Your mark:
 - Vim and command line test points ([0-10] each)
 - Points for presentations ([0-10] each):
 - [0-5] points from Laszlo + [0-5] points from students present, student points are averaged after discarding best and worst score

Course Programme - Pick Your Topic Now

1. Free operating systems, Debian, stable release
2. 1 Linux proficiency, terminal-based text editors, version control systems
2 Packaging with Autotools, Module::Build or distutils.core (Lothar)
3. Hypervisors, virtualization API, cloud computing platforms
4. 1 XML, XSD, BioXSD for interoperability in bioinformatics
2 Ontologies for interoperability in bioinformatics (*semester challenge*) (Matúš)
5. User management / directory services
6. 1 Mail, DNS
2 Web server
3 Databases and SQL, MySQL replication
7. 1 Web content management systems (Drupal and MediaWiki, Tim)
2 Non-Apache, non-MySQL (Manuel)
8. Network file systems, grid computing
9. 1 Packaging for major Linux distributions
2 Packaging for Debian and derivatives
10. Computer clusters and external services: excursion to LRZ
11. Discussion of semester challenge results



BioXSD

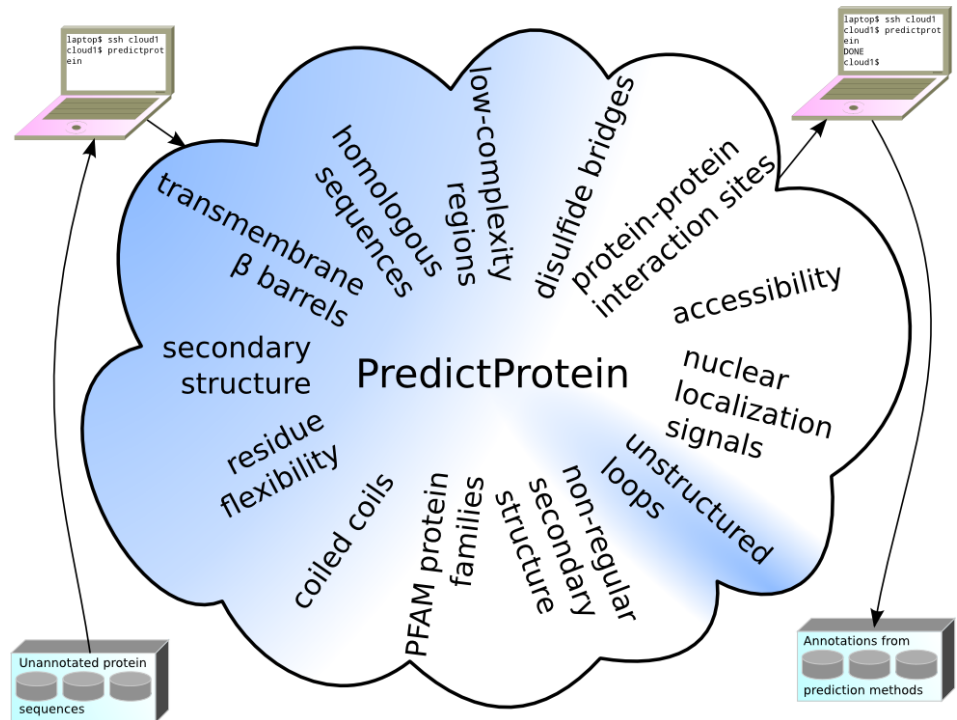


Semester Challenge

- Mandatory:
 - Prepare BioXSD-derived XSD schemas with semantic annotations for PredictProtein components
 - Prepare output converter programs using your XSD
- Optional:
 - Prepare WSDL file for PredictProtein component
 - Create and demonstrate operational web service, connect to other course members' web service
 - Use WSDL 2 and a RESTful web service

PredictProtein Components for XML Output Conversion

- disulfinder
- ncbi-seg
- ncoils
- norsnet
- norasp
- predictnls
- profbval
- profisis
- profphd
- proftmb
- A2M format (from blast, hmmer, hhblits, clustalw alignments)



Password

- Set your password now.
- i12r-tbl.informatik.tu-muenchen.de IN A
131.159.28.37
- SSH port: Garching ZIP code first four digits.
- Test your account.